

Building structured event indexes of large volumes of financial and economic data for decision making

EU - FP7 Work Programme Call FP7-ICT-2011-8-316404 Challenge 4.4 - Area Intelligent Information Management

> german.rigau@ehu.es http://adimen.si.ehu.es/~rigau http://www.newsreader-project.eu/



IXA NLP group http://ixa.si.ehu.es



- +
- OpeNER (2 years started on July 2012)
- NewsReader (3 years started on January 2013)
- LoCloud (3 years started on January 2013)
- Readers (3 years started on January 2013)

The NewsReader project

- Process financial news in 4 different European languages (English, Dutch, Spanish, Italian).
- Extract what happened to whom, when, and where.
 Align, storing provenance, not discarding any information.
- Distinguish unfolding story lines.
- Assist financial decision support by explaining current events.

- Automating Detective Work
- Discovering story lines on the Web

The NewsReader project

- Consortium
 - VU University Amsterdam (NL)
 - University of the Basque Country (ES)
 - Fondazione Bruno Kessler (IT)
 - LexisNexis (multinational)
 - ScraperWiki (GB)
 - Synerscope (NL)
- Funded by EU FP7 programme, grant 316404
- Jan. 2013 Dec. 2015



The problem

Thousands of decision makers

Daily documents considered by individual decision makers

Daily stream of documents

50-3,000

2 million sources



- → Duplications, repititions: new/old
- → Incomplete, piece-meal
- Inconsistent and contradictory
- → Coloured and opinionated
- Unauthorized

Archives covering decades

25 billion documents: news, company reports, manager biographies,...

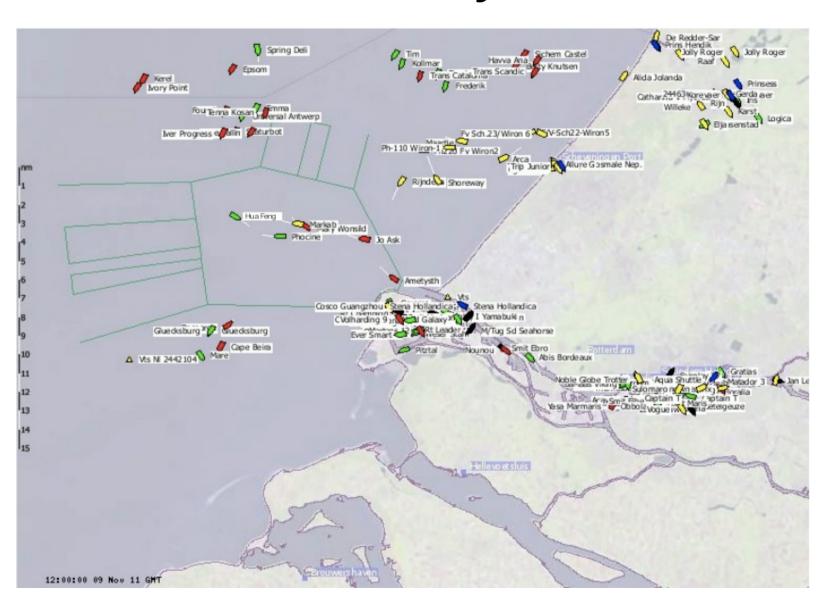
Events, sources and background data consulted Unknown volume







Example: the Hua Feng enters the harbor and nobody cares ...



http://shipphotos.wordpress.com



First, I want to say sorry for this very poor quality, but it was my lense, which was decentered in 2008. To the photo: Ex "Gulf Jash" (today "Hua Feng") passing Cuxhaven bound for Hamburg on the 07.10.2008.

http://www.tradewindsnews.com



Ship data from: Clarkson Research Services Limited

Gulf Jash (Ex: Probo Koala) Dead/Demolished 47,980 DWT, Bulk/Oil Carrier, Built 1989

Owners/Managers are Gulf Navigation, Built at Hanjin H.I., Double Hull, Panama Flagged, RINA Classed, Ice Strengthened C Class, Length Overall of 182.80 m., Length Between Perpendiculars of 175.72 m., Draught of 13.09 m., Beam of 31.98 m., 53.35 Tonnes per Centimetre Immersion, Gross Tonnage of 31,255, MAN B. & W. Engine, Speed of 15.00 kts at 28.00 tonnes per day, Heavy Fuel Oil, Horsepower of 12,800B at ...

Owner/Manager Details

Gulf Navigation Group

Specialised Details

Cargo Capacities of 51,000 cu.m. and 321,000 Barrels, Segregated Ballast Tanks, Clean Ballast Tanks, 7 Tanks, Epoxy Tank Coating, Grain Capacity of 50,669 cu.m., 7 Holds, 7 Hatches, Strengthened for Heavy Cargo, 2 Gantry Crane(s) with a safe working load of 28 tonnes.

Additional information

IDENTIFICATION: Exnames are Probo Koala, Probo Baoning. all Sign 3FOG9, IMO Number 8309816.

DIMENSIONS/TONNAGES: Moulded Depth of 10.03 m., Keel to mast air draft of 46.60 m., Tonnage of 26,146 Panama Canal Net, 26,837 Suez Canal Net, 14,257 International Net, 14,389 Light Displacement and 47,222 Dwt (long). ENGINE DETAILS: Engine Description 2 S.A. 4-cyl., Engine Model 4L80MC-E Mk1. CARGO HANDLING: 7 Centre Tanks with a capacity of 51,000 cu.m., Crude Oil Washing, Ore Cargo Capacity of 51,000 cu.m., Grab Fitted, Hatch Dimensions are 6 @ 13.50 x 20.20 m., 1 @ 13.50 x 12.20 m. SAFETY AND OTHER DETAILS: Last known special survey in July 2004, Clean Ballast Tanks, Inert Gas System.

Reported Sale and purchase information

Reported Sales available on Shipping Intelligence Network

Reported Fixture information

Reported Fixtures available on Shipping Intelligence Network



W

http://www.trafigura.com

Media Corrections

Times correction, 7 September 2009 (PDF)

"...Trafigura have always disputed that the dumping caused, or cou article and are defending the personal injury action being brought a the record straight."

BBC broadcast apology, 17 December 2009 (PDF)

"We accept that the independent experts who have considered this waste and these serious consequences. We withdraw those allega

"Our item headlined Success for the Guardian (26 April, page 2) erroneously linked the dumping of toxic waste in Ivory Coast from a vessel chartered by Trafigura with the deaths of a number of West Africans..."

Guardian

Independent apology, 27 February 2010 (Weblink)

"On 17 September 2009 our front page story 'Toxic shame' concerned the legal settlement in the UK of claims by individuals who alleged that they had been injured when waste from a Trafigura chartered ship was illegally dumped... The article stated that claimants had been maimed and wrongly suggested that, due to the settlement, claims of more serious injuries including miscarriages would not be tested in the High Court case. In fact such claims had already been withdrawn earlier last year..."

Times Online correction, 30 April 2010 (PDF)

"...We wish to make clear that the dumping was not carried out by Trafigura as the article may have suggested but by an independent local contractor without Trafigura's authority or knowledge. Furthermore, in September 2009 lawyers for Ivorians who were suing Trafigura over injuries allegedly caused by the dumping acknowledged that at worse the waste could only have caused flu-like symptoms..."

Guardian apology, 6 May 2010 (PDF)

"Our item headlined Success for the Guardian (26 April, page 2) erroneously linked the dumping of toxic waste in Ivory Coast from a vessel chartered by Trafigura with the deaths of a number of West Africans..."

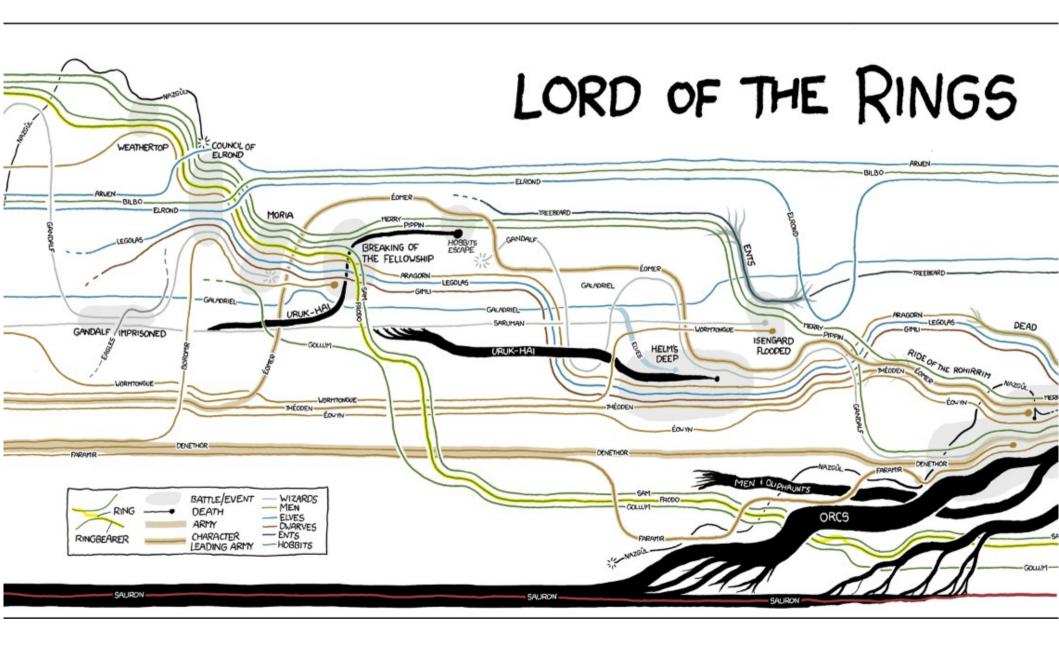
Economist correction, 16 December 2010 (Weblink)

"In our article on corporate leaks last week 'Be afraid', we said that in September 2009 WikiLeaks had posted a UN report showing Trafigura had dumped hazardous waste in Cote d'Ivoire. In fact, the report leaked was a draft of a study commissioned by Trafigura itself..."

The task

- discover the events that preceded and possibly led to a current event and summarize these for a human user
 - investigate participating actors, places, moments in time and a limited number of relations between these
 - follow all leads in parallel
 - search for information on the web
 - store the connections in an RDF graph, keeping precise track of the provenance of the connections
 - summarize and present the results as a story line

Story lines (= narrative charts?)





Building structured event indexes of large volumes of financial and economic data for decision making

EU - FP7 Work Programme Call FP7-ICT-2011-8-316404 Challenge 4.4 - Area Intelligent Information Management

> german.rigau@ehu.es http://adimen.si.ehu.es/~rigau http://www.newsreader-project.eu/

